Understanding context specificity

DE GRUYTER Diagnosis 2020; 7(3): 257–264

Abigail Konopasky*, Anthony R. Artino, Alexis Battista, Megan Ohmer, Paul A. Hemmer, Dario Torre, Divya Ramani, Jeroen van Merrienboer, Pim W. Teunissen, Elexis McBee, Temple Ratcliffe and Steven J. Durning

Understanding context specificity: the effect of contextual factors on clinical reasoning

https://doi.org/10.1515/dx-2020-0016
Received January 25, 2020; accepted March 11, 2020

Abstract

Background: Situated cognition theory argues that thinking is inextricably situated in a context. In clinical reasoning, this can lead to *context specificity*: a physician arriving at two different diagnoses for two patients with the same symptoms, findings, and diagnosis but different *contextual factors* (something beyond case content potentially influencing reasoning). This paper experimentally investigates the presence of and mechanisms behind context specificity by measuring differences in clinical reasoning performance in cases with and without contextual factors.

Methods: An experimental study was conducted in 2018–2019 with 39 resident and attending physicians in internal medicine. Participants viewed two outpatient clinic video cases (unstable angina and diabetes mellitus), one with distracting contextual factors and one without. After viewing each case, participants responded to six open-ended diagnostic items (e.g. problem list, leading diagnosis) and rated their cognitive load.

Results: Multivariate analysis of covariance (MANCOVA) results revealed significant differences in angina case performance with and without contextual factors [Pillai's trace = 0.72, F = 12.4, df = (6, 29), $p < 0.001$, $\eta_p^2 = 0.72$], with follow-up univariate analyses indicating that participants performed statistically significantly worse in cases with contextual factors on five of six items. There were no significant differences in diabetes cases between conditions. There was no statistically significant difference in cognitive load between conditions.

Conclusions: Using typical presentations of common diagnoses, and contextual factors typical for clinical practice, we provide ecologically valid evidence for the theoretically predicted negative effects of context specificity (i.e. for the angina case), with large effect sizes, offering insight into the persistence of diagnostic error.

Keywords: clinical reasoning; cognitive load; context; situated cognition.

***Corresponding author: Abigail Konopasky,** PhD, Assistant Professor of Medicine, The Henry M. Jackson Foundation for the Advancement of Military Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Rd, Bethesda, MD 20814, USA, Phone: +(301) 295-2904, E-mail: abigail.konopasky.ctr@usuhs.edu
Anthony R. Artino Jr.: Human Function, and Rehabilitation Sciences, School of Medicine and Health Sciences, The George Washington University, Washington, DC, USA, E-mail: aartino@email.gwu.edu
Alexis Battista: Department of Medicine, The Henry M. Jackson Foundation for the Advancement of Military Medicine, Uniformed Services University of the Health Sciences, Bethesda, USA, E-mail: alexis.battista.ctr@usuhs.edu
Megan Ohmer: Independent scholar, Bethesda, USA, E-mail: megandohmer@gmail.com
Paul A. Hemmer, Dario Torre, Elexis McBee and Steven J. Durning: Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA, E-mail: paul.hemmer@usuhs.edu (P.A. Hemmer); dario.torre@usuhs.edu (D. Torre); elexis.mcbee@usuhs.edu (E. McBee); steven.durning@usuhs.edu (S.J. Durning)
Divya Ramani: The Henry M. Jackson Foundation for the Advancement of Military Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA, E-mail: divya.ramani.ctr@usuhs.edu
Jeroen van Merrienboer and Pim W. Teunissen: School of Health Professions Education, Maastricht University, Maastricht, The Netherlands, E-mail: j.vanmerrienboer@maastrichtuniversity.nl (J. van Merrienboer); pwteunissen@gmail.com (P.W. Teunissen)
Temple Ratcliffe: Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA, E-mail: ratcliffe@uthscsa.edu

Introduction

Diagnostic error is a problem at the forefront of healthcare in the United States [1]. A recent National Academies of Science report concluded that diagnostic error is responsible for approximately 10% of patient deaths and between 6% and 17% of hospital adverse events [2]. Although diagnostic reasoning – and more broadly, clinical reasoning – is essential to patient care, our understanding of what influences it is limited, making it difficult to mitigate the effects of reasoning errors on patient care.

Clinical reasoning can be defined as the integration of clinical information, medical knowledge, and contextual

Published online May 02, 2020

factors to make patient care decisions [3–6]. A challenging phenomenon in clinical reasoning undoubtedly leading to diagnostic error is *context specificity*: when a physician arrives at two different diagnoses for two different patients who actually have the same symptoms, findings, and, ultimately, the same diagnosis [7, 8]. In a recent study of think-aloud reflections on reasoning, for instance, physicians note the presence of contextual factors (e.g. patient affect, diagnostic suggestion), which for some physicians seem to create uncertainty and difficulty with closure of the encounter [9]. We sought to build on this work by using a definition of context grounded in educational theory that could be applied to the rich complexity of practice settings in medicine. Doing so enabled us to empirically explore the phenomenon, one that is a source of unwanted variation in patient care. Thus, we defined case *context* as – going beyond case *content* – the individual (e.g. sleepiness, burnout, suggesting an incorrect diagnosis), physical (e.g. non-working electronic health record, allotted time for an appointment), and social aspects (e.g. challenging a physician's credentials) of a patient encounter to include the participants, the setting, and their interactions [10]. In this view, "context is not a fixed set of surrounding conditions, but a wider dynamical process of which the cognition of an individual is only a part" (p. xii) [11].

We approach the effects of context specificity on clinical reasoning using situated cognition theory, which argues that thinking (here, clinical reasoning) is inextricably bound within the context where it happens (complex interactions among patient, physician, and setting evolving over time) [10, 12]. This differs from traditional approaches to clinical reasoning that focus solely on the *content* of cases (e.g. symptoms, patient history, physical exam results) through mechanisms like illness scripts [13]. Instead, situated cognition recognises the importance of the participants, environment, and interactions therein, as noted in the above definition, offering a useful framework for understanding the effects of context specificity. Using this framework, we group contextual factors into those associated with the physician (e.g. fatigue), the patient (e.g. circuitous history), and the clinical environment (e.g. pressure to multitask [9, 14]; see Figure 1).

Another important concept is case specificity, which argues that different content, or different cases (e.g. diagnoses), can lead to different clinical reasoning performance. In explorations of context specificity, something more than case content (i.e. *case specificity* [15]) is influencing clinical reasoning: namely, what we refer to as contextual factors. While contextual factors can *positively* influence clinical reasoning, our work to date and the current study focus on understanding diagnostic error and factors that can *inhibit* it. Recent work suggests that contextual factors can inhibit clinical reasoning in both novice and more experienced physicians [7, 16]. This aligns with work in psychology suggesting that expert performance is not a stable *trait*, but a shifting, situation-based *state* [17–19].

Prior work suggests that one mechanism through which contextual factors may affect physicians is mental effort or *cognitive load*: constraints on how many information units one's working memory can hold and process at a time [7, 20, 21]. The assumption is that as a clinical encounter becomes more complex, with the introduction of different contextual factors, the associated cognitive load increases, potentially impairing clinical reasoning performance [7, 21]. Currently, however, relatively little is known about the relationship between cognitive load and clinical reasoning performance. Combining situated cognition's focus on the interactive elements of the clinical environment with cognitive load theory's focus on individual cognitive management of those elements offers

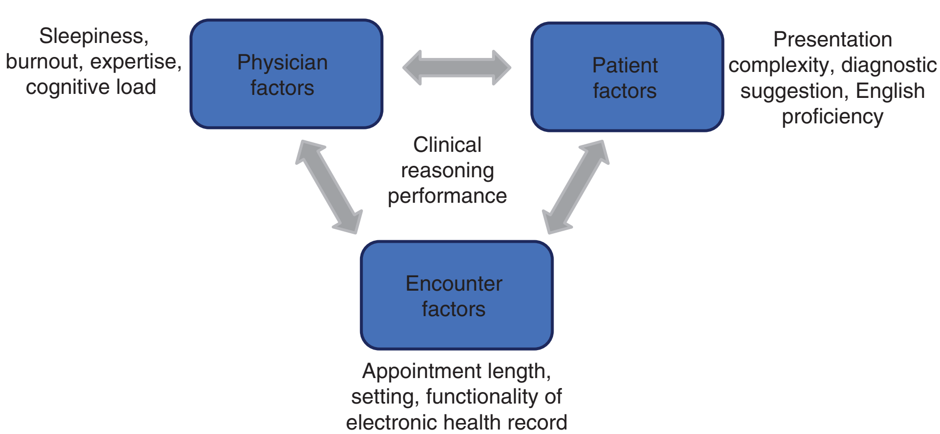

Figure 1: Contextual factors in clinical reasoning.

an opportunity to understand how different inhibiting contextual factors may influence clinical reasoning.

Context specificity is recognised as an important problem in medical education, and has been examined through the lens of prototype theory [22–24]. These earlier studies isolated particular elements of the case presentation like the language used, the timing of the presentation of a tentative differential diagnosis, and the familiarity of patient characteristics (e.g. name, age) and examined how they affected participants' (a) determination of a diagnosis and (b) identification of features of a case [23–31]. While this theoretical approach allowed them to tease apart how particular details of case presentations can affect the choice of leading diagnosis, it did not provide a framework for understanding the overall clinical situation in which this diagnostic choice takes place. Moreover, this work and recent work on contextual factors [14, 32] have looked only at the endpoint of diagnosis or feature identification, not the broader clinical reasoning *process* (e.g. evidence a physician offers for a given diagnosis) in which these decisions are grounded. Other recent work has taken up a socially situated theoretical model to better understand context specificity, but it has been exploratory and, as such, was conducted with no control group and relatively few participants [7, 14, 32–35].

Decrements in clinical reasoning can lead to unwanted variance in performance, patient morbidity and mortality, and/or excessive cost [2], so we must find innovative ways to examine and enhance clinical reasoning and context specificity more closely. Thus, the purpose of this study was to investigate both the presence of and the mechanisms behind context specificity, using contextual factors found to be important in our prior work [7, 9, 34]. Because clinical reasoning performance is related to years of professional experience [3] (our proxy for expertise), we also investigated whether experience affected performance. We also controlled for potential ordering effects. We asked:

– Is there a difference in clinical reasoning performance (as measured by open-ended diagnostic questions) when physicians diagnose cases with and without contextual factors?
– Is there a difference in self-rated cognitive load (i.e. mental effort) when physicians diagnose cases with and without contextual factors?

We hypothesised that participants would perform better and rate their mental effort as lower in diagnosing cases without inhibiting contextual factors. Further, based on the notion that expertise is a situation-based state (not an invariant trait), we hypothesised that contextual factors would affect physicians equally across experience levels. Finally, because participants regularly see far more than two cases daily, we hypothesised that there would be no ordering effects.

Materials and methods

To explore how context specificity may impair clinical reasoning, we designed two video simulation cases depicting patients with typical presentations of common diseases in primary care practice: new-onset diabetes mellitus and unstable angina. We believe that videos represent the optimal way to conduct this investigation as videos (widely used training tools) ensure all participants receive an identical "stimulus" to fully control both case content (identical content provided) and potentially relevant contextual factors (i.e. to empirically explore what may underpin context specificity). They lasted from just under 4 to 6.5 min and portrayed a clinical interview, a brief physical exam, and still screens of laboratory findings. We consulted with a group of internal medicine physicians to choose commonly encountered cases and contextual factors that are a part of everyday practice in internal medicine (and that tend to emerge with the types of cases we chose) to mitigate the effects of case (i.e. content) specificity [15] and enhance ecological validity. Prior to filming, the cases were screened by a panel of six medical education experts for authenticity and appropriateness (e.g. typicality).

Study participants were quasi-randomly assigned (based on their study day schedules) to one of two conditions: (a) diabetes case with inhibiting contextual factors (low English proficiency and a patient questioning the physician's credentials); angina case without contextual factors, or (b) angina case with inhibiting contextual factors (misleading diagnostic suggestion, patient reports history circuitously); diabetes case without contextual factors (see Figure 2). Furthermore, because this is early work in this area and because multiple contextual factors are typical across a busy day in practice, we used several contextual factors and two cases. The case content was controlled (i.e. identical presenting symptoms, language and gestures to represent those symptoms, and physical findings) so that it was the same for both diabetes and both angina cases; the only differences were the contextual factors. Also, conditions were balanced to control for potential ordering effects (i.e. whether the contextual factor case came first or second). After watching each case video, participants answered questions about diagnosis and mental effort.

Participants

A convenience sample (due to the study time demands and institutional requirements of volunteers only) of 39 resident and attending physicians in internal medicine were recruited from the Uniformed Services University of the Health Sciences, Walter Reed National Military Medical Center, and Naval Medical Center San Diego (see Table 1). We sought and received approval from all Institutional Review Boards (complying with the World Medical Association

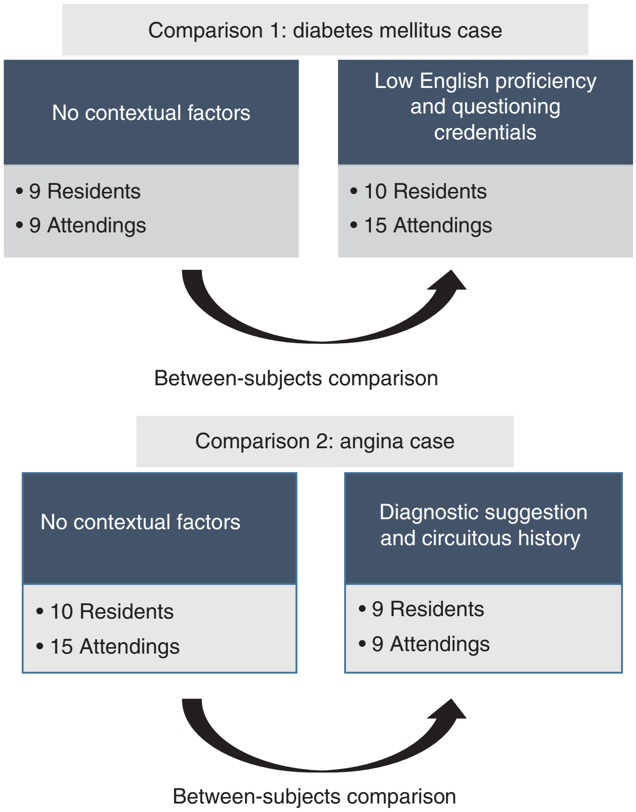

Figure 2: Study design.

Table 1: Participants' (12 females; 27 males) age and years in practice (n = 39).

	Mean (SD)	Range
Age, years	34 (6.8)	26–55
Years in practice	5.7 (6.4)	0–24

Declaration of Helsinki). Participants were allowed to take notes while watching the video cases.

Instruments

Post-encounter form: After viewing each video, participants completed a post-encounter form (PEF) asking for: (1) additional information they would like to obtain by history, (2) additional physical exam actions they would take, (3) a problem list, (4) a differential diagnosis, (5) a leading diagnosis, (6) supporting evidence for the diagnosis, and (7) management plans (not discussed in the current paper, which focuses on diagnostic reasoning; see Appendix 1 for survey). Participants were given up to 30 min (determined to be ample time in prior trials) for completing the items. Items were scored as in prior research where reliability and validity evidence for this instrument were established [33, 34]: each free-text response (most participants gave multiple responses for each question) was scored as correct (2 points), partially correct (1 point), or incorrect (0 points) based on a pre-determined scoring key developed by a panel of board certified internists and reported on in prior research [7, 33, 34] (with reliability between $\kappa = 0.82$ and $\kappa = 0.93$ in measure development). Participants were only able to give a single response for the leading diagnosis, but gave multiple responses for the other items. Three of the authors came to complete consensus on the scoring of all new utterances not on the key (less than 3% of responses). Then, in order to compare participants (who gave differing numbers of responses), a percentage correct score was calculated for each of the six items in this study, dividing total number of points received by total number of possible points (e.g. someone who gave five possible differential diagnoses has a total possible raw score of 10 for the differential diagnosis item).

Cognitive load: Cognitive load was measured through a single self-report item asking participants to: "Select your invested mental effort as you worked through the post-encounter form" [36]. The item, used in previous research [37–39], used a 10-point scale ranging from 1 (very low mental effort) to 10 (very high mental effort).

Data analysis

To determine if there were differences in clinical reasoning performance, two multivariate analyses of covariance (MANCOVAs) were performed with years of experience (i.e. number of years since medical school graduation) and case order (i.e. whether the diabetes or angina case was first) as covariates. The first between-subjects MANCOVA compared the six PEF scores on the diabetes case with and without contextual factors and the second compared those on the angina case with and without contextual factors. A power analysis indicated that a total sample size of 43 was needed to detect a large-sized effect at a significance level of 0.05, so the study was slightly underpowered.

To determine if there was a difference in self-reported cognitive load, two ANCOVAs were performed with experience and case order as the same covariates. The first between-subjects ANCOVA compared cognitive load scores on the angina case with and without contextual factors and the second compared cognitive load scores on the diabetes case with and without contextual factors.

Finally, to examine overall score differences, we averaged all PEF items for each case and conducted one-way ANCOVAs.

Results

Across all angina cases, the overall mean percentage scores on all PEF items ranged from 43% to 94% (m = 70%). The MANCOVA results revealed significant differences between the conditions with and without contextual factors [Pillai's trace = 0.72, $F = 12.4$, $df = (6, 29)$, $p < 0.001$, $\eta_p^2 = 0.72$], with no effects for case order or years of experience. Levene's test indicated equal variances for all dependent variables except the supporting evidence item ($F = 9.6$, $p = 0.006$). Follow-up univariate analyses indicated that participants performed significantly worse

Table 2: Univariate effects for MANCOVA analysis of angina case.

	No contextual factors (n = 23)		Contextual factors (n = 16)		F (6, 29)	p-Value	η_p^2
	Mean (%)	SD	Mean (%)	SD			
Additional history	0.72	0.15	0.63	0.17	3.8	0.06	0.10
Additional exam actions	0.79	0.18	0.50	0.21	24.1	0.000	0.42
Problem list	0.71	0.08	0.60	0.10	13.8	0.001	0.29
Differential diagnosis	0.68	0.13	0.56	0.16	6.7	0.014	0.17
Leading diagnosis	0.85	0.20	0.50	0.23	26.4	0.000	0.44
Supporting evidence	0.88	0.12	0.44	0.26	35.5	0.000	0.51

on all PEF items except additional history questions in the presence of a contextual factor, with large effect sizes [40] (see Table 2, showing the univariate effects). The ANCOVA of overall mean angina scores revealed significant differences between the conditions with and without contextual factors [$F = 82.7$, $df = (1, 34)$, $p < 0.001$, $\eta^2 = 0.71$], with a mean score of 77% without contextual factors [standard deviation (SD) = 0.07] and 55% with contextual factors (SD = 0.1).

Across all diabetes cases, the mean percentage scores on PEF items ranged from 37% to 84% (m = 70%). MANCOVA results revealed no significant differences between the conditions with and without contextual factors [Pillai's trace = 0.33, $F = 2.3$, $df = (6, 28)$, $p = 0.07$, $\eta_p^2 = 0.33$] or for either of the covariates. (See Table 3). The ANCOVA of overall mean diabetes PEF score showed no significant differences between conditions [$F = 0.1$, $df = (1, 33)$, $p = 0.76$, $\eta_p^2 = 0.003$].

There was no significant difference in self-rated cognitive load with or without contextual factors for the diabetes [$F (1, 40) = 6.1$, $p = 0.38$] or angina [$F (1, 40) = 1.2$, $p = 0.52$] cases. Levene's test indicated equal variances for both diabetes ($F = 0.61$, $p = 0.61$) and angina ($F = 0.55$, $p = 0.65$). We did, however, observe a trend in cognitive load: self-reported cognitive load was higher in the presence of contextual factors for both diabetes (m = 4.85 for no contextual factors and m = 5.72 for contextual factors) and angina (m = 6.64 for no contextual factors and m = 6.95 for contextual factors).

Discussion

In this investigation, we sought to experimentally test the phenomenon of context specificity in a group of experienced internists using a theoretically grounded approach. While context specificity in clinical reasoning has become an important area of study [7–9, 26, 29, 31, 34, 41–43], to the best of our knowledge, this is the first study to use a robust socially situated theoretical framework, carefully controlled stimulus, ecologically valid measure of clinical reasoning, and fairly large sample of participants. Using typical presentations of common diagnoses, and common contextual factors, our results suggest that the negative effects of context specificity may in part be dependent on case content, i.e. context specificity holds in our data for the angina, but not diabetes, case (with an aggregated PEF score and with five of six individual PEF measures). While prior context specificity work demonstrated significant correlations between PEF items and cognitive load, those studies did not include control groups *without* contextual factors [9, 34]. This study extends that work, demonstrating significant performance differences in an angina case across five domains: additional exam actions,

Table 3: Descriptive statistics for MANCOVA analysis of diabetes case.

	No contextual factors (n = 23)		Contextual factors (n = 16)	
	Mean (%)	SD	Mean (%)	SD
Additional history	0.51	0.21	0.62	0.15
Additional exam actions	0.78	0.16	0.77	0.19
Problem list	0.74	0.14	0.75	0.08
Differential diagnosis	0.50	0.11	0.50	0.14
Leading diagnosis	0.78	0.19	0.74	0.23
Supporting evidence	0.89	0.07	0.81	0.21

problem list, differential diagnosis, leading diagnosis, and supporting evidence. Moreover, this effect held *across* years of experience, indicating the importance of deliberate practice [17, 19] and context in physicians' continuing education: understanding case content is not adequate – physicians must carefully practice reasoning with that content across environments.

Although we found context specificity for the angina case, we did not find such an effect for the diabetes case, despite similar score ranges across the cases and our experts' judgement that both cases were equally common and typical. This finding could be due to inadequate power or to the dose or quality of the contextual factors in the two cases (chosen for their ecological validity with respect to the content area); i.e. perhaps circuitous history poses more or different challenges to clinical reasoning than the agitated non-native English speaker. This could also be evidence of case specificity: some aspect of the content of the diabetes case could make it easier to circumvent contextual effects when compared to the angina case, which has greater acuity and fewer specific diagnostic signs. Thus, neither case content nor contextual factor is the *sole* predictor of clinical reasoning performance. As others have argued, there is no single cause for diagnostic errors, but a nuanced and complex system of interacting conditions, including context *and* content [15, 44].

Regarding self-reported cognitive load, while the scores trended in the expected direction for both cases (higher cognitive load with contextual factors), these trends were not statistically significant. This could be due to inadequate power or to problems with our single-item measure of cognitive load (i.e. inadequate sensitivity). This could also indicate that the cause of the performance decline in the angina case is not the result of increased cognitive load, but due to some other set of factors, such as different emotional reactions [45]. Future work in this area could elucidate this.

As the expertise literature would predict, controlling for years of experience did not eliminate the effect of contextual factors [18, 19]. In other words, context specificity effects are not limited to newer physicians. Thus, future support tools should be developed not just for residents, but for attending physicians as well. We did note, however, that more experienced physicians performed significantly better on the additional history and additional exam items. Perhaps these diagnostic tasks become more automated or scripted over the years than others (e.g. leading diagnoses). Again, this resonates with prior work indicating that tasks *within* clinical reasoning are equally as important as the broader content or context [15, 44, 46].

There are several important study limitations. First, the sample size of 39 participants is only 31% women and, moreover, is relatively small for the statistical test. Yet, this ratio is representative for the participants' institutions and this is actually a large number of participants for a clinical reasoning study (one requiring 2 h of physician participant time). Second, we only ran two sets of cases, using different contextual factors in each (diabetes and angina), potentially affecting the interpretation (i.e. effects could be the result of a *specific type* of contextual factor). Yet, based on our prior work, we wanted to include various contextual factors educators and researchers have hypothesised to be important.

This study has several important practical implications. First, and perhaps most centrally for training and continuing education, these results indicate the importance of identifying areas *other* than content or medical knowledge that contribute to establishing diagnoses as the findings were demonstrated in both residents and attendings. Explicit education in cases with identified contextual factors – perhaps through simulation – could potentially mitigate the negative effects of context specificity. Future work might explore different contextual factors (e.g. appointment length) and their mitigation or elimination. Additionally, the theoretical model and proposed contextual factors could be explored in more authentic settings such as the clinics or wards to better understand context specificity.

Second, this study suggests that we may be underestimating the effect of case specificity on error. Diabetes and angina are two common content areas. While case specificity may explain less variance than individual items within cases in some contexts [15], it appears to have a significant effect on those items when certain contextual factors are introduced.

Third, while serving as an optimal platform for exploring context specificity, the videos themselves may have induced additional cognitive load. We should not assume that clinical reasoning "in person" is the same as clinical reasoning mediated through technology. As such, our findings may have implications for clinical reasoning in technology-enhanced contexts, such as telehealth.

In conclusion, our findings are consistent with expectations of situated cognition theory, which would predict that interactions among different contextual factors and different case content would engender different clinical reasoning challenges. Diagnostic error plagues our healthcare system, and we believe that a work like this can help illuminate the vexing phenomenon of context specificity. Additionally, this work points to the need for interventions to reduce unwanted performance variance. Such interventions could benefit healthcare systems nationwide.

Appendix 1

Post-encounter form survey

Q1) What else do you want to ask this patient? List one to five questions.

Q2) What else would you want to look for on this patient's physical exam? (List one to five items).

Q3) Write a complete problem list.

Q4) What is your differential diagnosis? (Please list in order of likelihood and list at least three responses).

Q5) What is your leading diagnosis?

Q6) What data supports this diagnosis? (List one to five pieces of evidence).

Q7) What is your treatment plan for this patient (diagnostic and/or therapeutic)?

Author contributions: All authors collaborated together on the research design and data collection. Konopasky ran the analyses and wrote up the Results section. Konopasky and the remaining authors co-wrote the remainder of the paper. All authors offered substantive revisions and approved of the final version of the paper. All authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: This study was supported by a grant from the JPC-1, CDMRP – Congressionally Directed Medical Research Program [Funder Id: http://dx.doi.org/10.13039/100000090 (#NH83382416)]. This work was made possible by funding from the Congressionally Directed Medical Research Program, Joint Program Committee 1.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organisation(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.
Disclaimer: The opinions and assertions expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University or the Department of Defense.

References

1. Singh H, Graber ML. Improving diagnosis in health care – the next imperative for patient safety. N Engl J Med 2015;373:2493.
2. National Academies of Sciences and Medicine E. Improving diagnosis in health care. Washington, DC: National Academies Press, 2016.
3. Croskerry P. A universal model of diagnostic reasoning. Acad Med 2009;84:1022–8.
4. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving an analysis of clinical reasoning. Cambridge, MA: Harvard University Press, 1978.
5. Heneghan C, Glasziou P, Thompson M, Rose P, Balla J, Lasserson D, et al. Diagnostic strategies used in primary care. Br Med J 2009;338(apr20_1):b946.
6. Young M, Thomas A, Lubarsky S, Ballard T, Gordon D, Gruppen LD, et al. Drawing boundaries: the difficulty in defining clinical reasoning. Acad Med 2018;93:990–5.
7. Durning SJ, Artino AR, Boulet JR, Dorrance K, van der Vleuten C, Schuwirth L. The impact of selected contextual factors on experts' clinical reasoning performance (does context impact clinical reasoning performance in experts?). Adv Heal Sci Educ 2012;17:65–79.
8. Eva KW. What every teacher needs to know about clinical reasoning. Med Educ 2005;39:98–106.
9. McBee E, Ratcliffe T, Picho K, Artino AR, Schuwirth L, Kelly W, et al. Consequences of contextual factors on clinical reasoning in resident physicians. Adv Heal Sci Educ 2015;20:1225–36.
10. Durning SJ, Artino Jr AR, Pangaro LN, van der Vleuten C, Schuwirth L. Perspective: redefining context in the clinical encounter: implications for research and training in medical education. Acad Med 2010;85:894–901.
11. Hutchins E. Cognition in the wild. Cambridge, MA: MIT Press, 1995.
12. Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. Educ Res 1989;18:32–42.
13. Custers EJ. Thirty years of illness scripts: theoretical origins and practical applications. Med Teach 2015;37:457–62.
14. Westbrook JI, Raban MZ, Walter SR, Douglas H. Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: a prospective, direct observation study. BMJ Qual Saf 2018;27:655–63.
15. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? Med Educ 2006;40:618–23.
16. Durning S, Artino Jr AR, Pangaro L, van der Vleuten CP, Schuwirth L. Context and clinical reasoning: understanding the perspective of the expert's voice. Med Educ 2011;45:927–38.
17. Chow DL, Miller SD, Seidel JA, Kane RT, Thornton JA, Andrews WP. The role of deliberate practice in the development of highly effective psychotherapists. Psychotherapy 2015;52:337–45.
18. Ericsson KA. The influence of experience and deliberate practice on the development of superior expert performance. Cambridge Handb Expert Expert Perform 2006;38:685–705.
19. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. Psychol Rev 1993;100:363.

20. Paas F, Tuovinen JE, Tabbers H, Van Gerven PW. Cognitive load measurement as a means to advance cognitive load theory. Educ Psychol 2003;38:63–71.
21. Young JQ, Van Merrienboer J, Durning S, Ten Cate O. Cognitive load theory: implications for medical education: AMEE guide no 86. Med Teach 2014;36:371–84.
22. Norman GR, Brooks LR. The non-analytical basis of clinical reasoning. Adv Heal Sci Educ 1997;2:173–84.
23. Hatala R, Norman GR, Brooks LR. Influence of a single example on subsequent electrocardiogram interpretation. Teach Learn Med 1999;11:110–7.
24. Kulatunga-Moruzi C, Brooks LR, Norman GR. Using comprehensive feature lists to bias medical diagnosis. J Exp Psychol Learn Mem Cogn 2004;30:563.
25. Hatala R, Norman GR, Brooks LR. Impact of a clinical scenario on accuracy of electrocardiogram interpretation. J Gen Intern Med 1999;14:126–9.
26. Brooks LR, LeBlanc VR, Norman GR. On the difficulty of noticing obvious features in patient appearance. Psychol Sci 2000;11:112–7.
27. Norman GR, Brooks LR, Colle CL, Hatala RM. The benefit of diagnostic hypotheses in clinical reasoning: experimental study of an instructional intervention for forward and backward reasoning. Cogn Instr 1999;17:433–48.
28. LeBlanc VR, Norman GR, Brooks LR. Effect of a diagnostic suggestion on diagnostic accuracy and identification of clinical features. Acad Med 2001;76:S18–20.
29. LeBlanc VR, Brooks LR, Norman GR. Believing is seeing: the influence of a diagnostic hypothesis on the interpretation of clinical features. Acad Med 2002;77:S67–9.
30. Young M, Brooks L, Norman G. Found in translation: the impact of familiar symptom descriptions on diagnosis in novices. Med Educ 2007;41:1146–51.
31. Young ME, Brooks LR, Norman GR. The influence of familiar non-diagnostic information on the diagnostic decisions of novices. Med Educ 2011;45:407–14.
32. Mercuri M, Sherbino J, Sedran RJ, Frank JR, Gafni A, Norman G. When guidelines don't guide: the effect of patient context on management decisions based on clinical practice guidelines. Acad Med 2015;90:191–6.
33. Durning SJ, Artino A, Boulet J, La Rochelle J, Van Der Vleuten C, et al. The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. Med Teach 2012;34:30–7.
34. McBee E, Ratcliffe T, Picho K, Schuwirth L, Artino AR, Yepes-Rios AM, et al. Contextual factors and clinical reasoning: differences in diagnostic and therapeutic reasoning in board certified versus resident physicians. BMC Med Educ 2017;17:211.
35. ten Cate O, Durning SJ. Understanding clinical reasoning from multiple perspectives: a conceptual and theoretical overview. Principles and Practice of Case-based Clinical Reasoning Education 2018:35–46.
36. Brunken R, Seufert T, Paas F. Measuring cognitive load. In: Paas J, Moreno R, Brunken R, editors. Cognitive load theory. London: Cambridge University Press, 2010:181–202.
37. Szulewski A, Gegenfurtner A, Howes DW, Sivilotti ML, van Merriënboer JJ. Measuring physician cognitive load: validity evidence for a physiologic and a psychometric tool. Adv Heal Sci Educ 2017;22:951–68.
38. Ratcliffe TA, McBee E, Schuwirth L, Picho K, Van der Vleuten C, Artino A, et al. Exploring implications of context specificity and cognitive load in residents. MedEdPublish 2017;6.
39. Young JQ, Irby DM, Barilla-LaBarca M-L, ten Cate O, O'Sullivan PS. Measuring cognitive load: mixed results from a handover simulation for medical students. Perspect Med Educ 2016;5:24–32.
40. Richardson JT. Eta squared and partial eta squared as measures of effect size in educational research. Educ Res Rev 2011;6:135–47.
41. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. N Engl J Med 2006;355:2217–25.
42. Cook DA, Sherbino J, Durning SJ. Management reasoning: beyond the diagnosis. J Am Med Assoc 2018;319:2267–68.
43. Artino Jr AR, Cleary TJ, Dong T, Hemmer PA, Durning SJ. Exploring clinical reasoning in novices: a self-regulated learning microanalytic assessment approach. Med Educ 2014;48:280–91.
44. Wimmers PF, Fung C. The impact of case specificity and generalisable skills on clinical performance: a correlated traits-correlated methods approach. Med Educ 2008;42:580–88.
45. Croskerry P, Abbass A, Wu AW. Emotional influences in patient safety. J Patient Saf 2010;6:199–205.
46. Juma S, Goldszmidt M. What physicians reason about during admission case review. Adv Heal Sci Educ 2017;22:691–711.